

Northumbria Research Link

Citation: Gunenc, Cevher, O'Shea, Laura E. and Dickens, Geoffrey (2018) Structured risk assessment for reduction of multiple risk outcomes in a secure mental health setting: Use of the START. *Criminal Behaviour and Mental Health*, 28 (1). pp. 61-71. ISSN 0957-9664

Published by: Wiley

URL: <https://doi.org/10.1002/cbm.2036> <<https://doi.org/10.1002/cbm.2036>>

This version was downloaded from Northumbria Research Link:
<http://nrl.northumbria.ac.uk/id/eprint/44511/>

Northumbria University has developed Northumbria Research Link (NRL) to enable users to access the University's research output. Copyright © and moral rights for items on NRL are retained by the individual author(s) and/or other copyright owners. Single copies of full items can be reproduced, displayed or performed, and given to third parties in any format or medium for personal research or study, educational, or not-for-profit purposes without prior permission or charge, provided the authors, title and full bibliographic details are given, as well as a hyperlink and/or URL to the original metadata page. The content must not be changed in any way. Full items must not be sold commercially in any format or medium without formal permission of the copyright holder. The full policy is available online: <http://nrl.northumbria.ac.uk/policies.html>

This document may differ from the final, published version of the research and has been made available online in accordance with publisher policies. To read and/or cite from the published version of the research, please visit the publisher's website (a subscription may be required.)



**Northumbria
University**
NEWCASTLE



UniversityLibrary

Structured risk assessment for reduction of multiple risk outcomes in a secure mental health setting: use of the START

Cevher Gunenc, Laura E. O'Shea and Geoffrey L. Dickens

This is the peer reviewed version of the following article:

Gunenc, C., O'Shea, L.E. and Dickens, G.L. 2017. Structured risk assessment for reduction of multiple risk outcomes in a secure mental health setting: use of the START. *Criminal Behaviour and Mental Health*.

which has been published in final form at doi:

<http://dx.doi.org/10.1002/cbm.2036>

This article may be used for non-commercial purposes in accordance with the [Wiley Terms and Conditions for Self-Archiving](#)

Structured risk assessment for reduction of multiple risk outcomes in a secure mental health setting: Use of the START.

Short title: START for risk outcome reduction

Cevher Gunenc, Plymouth University

Laura E. O'Shea, St Andrew's Healthcare

Geoffrey L. Dickens*, Abertay University and NHS Fife

*** Corresponding author**

Email: g.dickens@abertay.ac.uk

Telephone: 07914157365

Address: Division of Mental Health Nursing and Counselling, Abertay University, Bell Street, Dundee, DD1 1HG. United Kingdom.

Abstract

Background: Structured risk assessment to aid prediction and prevention of risk behaviours in secure settings is common; the expected benefits have rarely been investigated.

Aims: To determine whether adverse outcomes (physical and verbal aggression, self-harm, victimisation, self-neglect, unauthorised leave, substance abuse) reduced after patients were assessed with the Short-Term Assessment of Risk and Treatability (START).

Methods: In a naturalistic study, anonymised demographic and clinical information were collected from 50 male patients. Data included START assessment and frequency of target behaviour data for three months before and after first assessment. Chi square and linear mixed models analyses were used to determine whether there was any change in the behaviours of interest.

Results: There were no significant changes in physical or verbal aggression over time, although a tendency towards fewer incidents was apparent. Other adverse behaviours were very infrequent at baseline, precluding adequate analysis.

Conclusions: In this small sample, START risk assessment did not achieve its primary purpose of significant reduction in adverse behaviours. Although we based our sample size on a power calculation, we may have over-estimated the size of anticipated change. Further, the three-month comparison periods before and after the assessment follow-up period, were rather short. Accordingly, we recommend more research on the value of this tool in practice rather than discontinuing its use.

Keywords: START, risk assessment, risk management, violence, aggression, self-harm

Introduction

Violence, aggression, and other harmful behaviours, are common among secure psychiatric inpatients (Foster et al., 2007), and the use of structured tools to aid risk assessment is common practice (O'Shea et al., 2014a). There is substantial evidence that some tools (e.g., HCR-20; Webster et al., 1997) help to improve violence prediction (O'Shea & Dickens, 2014a), and some evidence about prediction of self-harm (O'Shea et al., 2014b), verbal aggression (Gunenc et al., 2015), and sexually inappropriate behaviour (O'Shea et al., 2015). While accurate prediction of risk behaviours is a key aim of practice, the ultimate goal of risk assessment is to facilitate successful management and prevention or reduction of harm (Hart, 1998). Studies about the effectiveness of assessment as an intervention in secure inpatient settings are, however, scarce (Large & Nielssen, 2012) and, for non-violent outcomes, absent. The generalisability of results from trials of standardised risk of violence assessments has been limited due to differences in baseline characteristics of control and experimental subjects, lack of blinding of clinical staff to study allocation, and small sample size (Abderhalden et al., 2008; van de Sande, 2011). In addition, risk flagging systems which simply identify the presence of previous risk behaviour, as opposed to structured risk assessments which aim to identify empirical risk factors, have shown some promise. In one study (Kling et al., 2011), violent incidents reduced during use of the flagging system, but returned to pre-assessment levels after the trial, suggesting that risk flagging offers limited information about intervention targets (Douglas & Skeem, 2005). In contrast, assessments based on dynamic risk factors such as impulsivity, medication non-adherence, and lack of social skills, have the potential to reduce risk outcomes because they facilitate identification of modifiable treatment targets (Hanson & Harris, 2000).

The Short-Term Assessment of Risk and Treatability (START; Webster et al., 2009) is predicated on identification of dynamic risk factors and the prediction of multiple adverse outcomes: violence, self-harm, suicide, self-neglect, victimisation, substance abuse, and unauthorised leave. A further advantage of the tool is its additional incorporation of protective factors, or strengths. Thus far, only Troquete et al. (2013) have examined change

in adverse event incidence following START assessment in a community setting, and then only for violent outcomes. The number of violent and criminal behaviours that had taken place in each study arm in the 6-month follow-up period following intervention (START administration plus shared care planning approach or treatment as usual) fell significantly compared to baseline, but between-group difference was not significant. Therefore, there is a lack of evidence about the effectiveness of START as an intervention for reducing risk behaviours, especially for its full range of intended assessment targets and in the secure inpatient setting.

The primary aim of our study was to examine change in the incidence of risk behaviours in a secure inpatient psychiatric setting where use of the START had been mandated. We hypothesised that there would be a reduction in risk incidents, and in the proportion of patients engaging in each type of risk behaviour, following risk assessment for the START outcomes. Secondary aims were to compare changes between patient groups defined by different demographic or clinical features and to inform decisions about the sample size necessary for a full prospective trial to detect significant change in less well-studied non-aggressive but adverse behaviours following risk assessment.

Method

Setting and participants

St. Andrew's is a UK charity providing specialist locked, low-, and medium-secure inpatient care for people with mental disorder. Eligible patients were consecutive admissions between May 2011 and March 2012 to the male, adult, mental illness/personality disorder services who had been assessed using the START for the first time since its introduction into the service. Exclusions were those solely with an intellectual disability diagnosis. Patients must have been present in the service continuously for three months before and after START assessment. Sample size was predicated on a power calculation using a dichotomous 'any

aggression' variable which in a previous study in the current setting was 62% over a similar 3-month follow-up period (O'Shea et al., 2016). In anticipation of an 80% chance ($\beta=0.8$) of detecting a 50% reduction following risk assessment ($\alpha=0.05$) a sample of 37 in each of the pre- and post- assessment conditions was required (Sealed Envelope Ltd., 2012). A reduction of 50% would, however, represent a very large effect size ($d=1.1$). An increase in the sample size to 50 brought the required effect size required into the medium range ($d=0.68$; Lipsey & Wilson, 2001).

Design and procedure

The study followed a naturalistic, pseudo-prospective AB design using data generated in routine clinical practice. Since all data were anonymised we could not seek individual informed consent and NHS Research Ethics Committee approval was not required. The study was approved by the St Andrew's clinical audit manager as a service evaluation.

Risk assessment

The START (Webster et al., 2009) is a structured professional judgement tool comprising 20 empirically-derived dynamic items. Each is rated twice on a 3-point scale: once for vulnerabilities (risks) and once for strengths (protective factors) (0 = no/minimal vulnerability or strength; 1 = moderate vulnerability or strength; 2 = high vulnerability or strength). Based on item ratings, historical and clinical information, and information about the nature, imminence, and severity of risks, raters formulate a specific risk estimate (SRE: 'low', 'moderate', or 'high') for each of the seven risk outcomes. The START should be completed by a “number of mental health specialists who work together as a team” (Webster et al., 2009: p.24) through a discursive, consensual process in order to incorporate a range of professional opinions for each case. The START manual makes no recommendations about the number, training, or qualifications of raters. In the setting for this study, raters attend a structured 1-day training based on the START manual and cover theoretical and practical aspects of the tool. Training includes teaching, team discussion and rating of case stories derived from real patients, but with any identifying features removed as well as feedback, and

further discussion in the context of ratings given by START experts, and those of other teams at previous training. It is required that the completed START for each patient is signed off by three multidisciplinary clinical team members from different professions (psychiatrist, psychologist, nurse, occupational therapist or social worker). Rating is completed every 3-months and this is supported by regular audit of a sample of cases to ensure maintenance of skills. Since START assessments included in this study were rated for clinical purposes, we were unable to calculate inter-rater reliability statistics; a strategy which has been applied in previous research (e.g., Inett et al., 2014; Nonstad et al., 2010; Quinn et al., 2013). Total scores on the strength and vulnerability scales were prorated to account for missing items, in accordance with guidelines in the START manual (Webster et al., 2009).

Demographic and clinical data

Anonymised demographic information concerning patients', age at assessment, ethnicity, ICD-10 (World Health Organization, 1992) diagnoses, admission date, security level, and legal status was retrieved.

Risk outcomes

Local policy mandates that a free-text electronic progress note is entered by a qualified member of the clinical team for each patient on each shift, thus 2-3 times per day. At the time of entry, staff electronically flag the note if any of a range of risk outcomes has occurred; incidents can be flagged for multiple outcomes. We collated incidents flagged as "aggression – physical", "aggression – verbal", "absconding", "self-harm/suicide", "self-neglect" "substance misuse", and "vulnerability" that occurred in the 3-month period before and after START assessment for each patient. Criteria from the START Outcome Scale (SOS; Nicholls et al., 2007) were used to code incidents in the categories physical aggression, verbal aggression, self-harm/suicide behaviours, self-neglect, unauthorised leave, substance abuse, and victimisation. A sample of data for each outcome was coded independently by two authors (LO, CG) to determine inter-rater reliability, which was in the good to perfect range (Kappa 0.64 – 1.00). The SOS includes a total of 12 outcome categories: verbal aggression,

aggression against property, physical aggression against others, sexual aggression, self-harm, suicide ideation and planning, suicidal behaviour, self-neglect, unauthorised leave, substance abuse, being victimised and stalking. Each is rated on a 5-point scale (0 absent; 1 mild severity through to 4 - most severe). Incidents of sexual aggression and stalking were not coded unless flagged as appropriate since the START is not used to predict these outcomes. Further, aggression against property was included in the aggression against others category, and self-harm, suicide ideation, and suicidal behaviour were coded under one heading (self-harm) due to the difficulty in determining intent (Gray et al., 2011). Raters judged whether each flagged note met the requirement for a level 1 incident or above for any of the outcome categories examined.

Data analysis

Statistical analysis was performed using SPSS version 18 for Windows (SPSS, 2009). Descriptive statistics were calculated for sample characteristics, mean START strength and vulnerability scores of the sample, the proportion of patients with low, moderate, and high SREs for each risk outcome, the proportion of patients who engaged in each outcome and the number of incidents of each risk outcome three months before and after the first START assessment. A linear mixed models analysis was used to examine baseline number of incidents before assessment and changes in number of incidents of all risk outcomes after assessment. Use of mixed models allowed for analysis of differences in number of incidents at baseline, and rates of change between multiple patient groups relative to the reference category. Patients were compared in terms of: 1) ethnicity; Non-White and those of unknown ethnicity were compared to White patients (the reference category), 2) diagnosis; personality disorder, developmental disorder, and other/none/multiple diagnoses were compared to schizophrenia (the reference category), 3) ward security level; low secure was compared to medium secure (the reference category), 4) Mental Health Act 1983, amended 2007 section; forensic was compared to civil (the reference category), and 5) age: 35 or over was compared to 34 and under (the reference category). Changes in the proportion of patients rated as being within a particular risk category before and after risk assessment were analysed using the chi

square test (physical aggression, verbal aggression, self-harm, victimisation) and Fisher's exact test (self-neglect, unauthorised leave, substance abuse) according to cell size.

Results

Participants

Sixty-seven male inpatients met the clinical inclusion criteria; two were excluded due to excessive missing information from their START assessment, nine due to not being present in the service for three months before assessment, and six because they were not present for three months following assessment, leaving a sample of 50. Twenty-four (48%) patients resided on low secure and locked wards and 26 (52%) on medium secure wards. $M_{Age}=34.28$ years ($SD=15.02$; range 18-72). The sample comprised three (6%) Asian patients, three (6%) Black patients, 19 (38%) White patients, and 25 (50%) patients with unspecified ethnicity. Seventeen (34%) patients were detained under civil sections and 33 (66%) under forensic sections of the Mental Health Act 1983, amended 2007. Mean (M) admission to assessment time was 269.98 ($SD=105.28$) days. **Tewenty eight (56%) patients had a single diagnosis and** the remainder two or more diagnoses. The two largest diagnostic groups, which were more or less mutually exclusive, were schizophrenia, schizotypal and delusional disorders (24, 48%) and disorders of psychological development (23, 46%); other, generally co-morbid diagnoses included organic brain disorders (6, 12%).

START scores

The mean strength score was 17.55 ($SD=7.64$) and the mean vulnerability score 24.22($SD=7.09$). Table 1 shows the frequency distribution of patients who were estimated to have low, moderate, and high risk of each of the seven risk outcomes.

[INSERT TABLE 1 ABOUT HERE]

Number of incidents

Table 2 shows that the total number of incidents of physical aggression, verbal aggression, self-harm, victimisation, and substance abuse was lower three months after assessment than three months before assessment. Incidents of self-neglect and unauthorised leave, however, were slightly higher. None of the differences were statistically significant.

[INSERT TABLE 2 ABOUT HERE]

Table 3 shows that changes in the proportion of patients involved in incidents were not statistically significant.

[INSERT TABLE 3 ABOUT HERE]

Differences in between group change over time

. There were only three significant between group differences in change in behaviour. Incidents of verbal aggression by patients of unknown ethnicity reduced by an additional 4.88 incidents after assessment compared to White patients ($t[50] = -2.69, p = 0.010$). Verbal aggression reduced by an additional 6.27 incidents in those with developmental disorder ($t[50] = -2.47, p = .017$) and 6.95 incidents in those with other/none/multiple diagnoses ($t[50] = -3.35, p = .002$), compared to those with schizophrenia.

Discussion

We detected a no significant reduction in the number of physically or verbally aggressive acts, or other acts of victimisation against others. Nor was there any reduction in self-harm or substance abuse incidents in the three months after START assessment. Base rates for incidents of self-neglect or unauthorised leave were very low, reducing the potential

for finding significant change. Although the proportion of individuals who did something aggressive during the period showed some reduction, this was not statistically significant.

These results are inconsistent with previous research. Troquete et al. (2013) found a significant decrease in the number of incidents after START assessment compared to baseline number of incidents during a 6-month follow-up period, but not compared to a treatment as usual control group. The longer follow-up period used may partially explain these findings. Troquete et al. (2013) believe that the negative aspect of their finding could be due to the already high standard of care, an argument that could be deployed about the findings in this study. While this may hold true for relatively high base-rate violent outcomes, it is more likely that low base rates outcomes require significantly larger sample sizes if meaningful change is to be detected. In this study, for example, the base rate of self-neglect was 6%; to detect a 50% reduction following START assessment would require a sample of $N=746$ at pre- and at post- assessment (Sealed Envelope Ltd., 2012); to detect reductions in base rate for unauthorised leave from 2% to 1% would require in excess of 2,300 patients at both points. These requirements render some aspects of the START difficult to test empirically in terms of their effectiveness as an intervention. Another hypothesis that could be derived from these findings is that it is the iterative process of assessment, management plan adjustment, and re-assessment that has the potential to reduce risk behaviours rather than a single intervention. Rather in support of this possibility, most of the patients were assigned an SRE of 'low' for all of the risk outcomes excluding suicidal behaviours. This indicates a possibility that the interventions provided to these patients may have already successfully reduced risk behaviours and, in turn, the potential for further reductions in the number of incidents. Testing this proposition would require access to a population naive to structured risk assessment.

Although there were no overall changes in number of incidents or proportions of patients with the difficulties measures, there was a suggestion that management of some diagnostic groups may be disproportionately helped by START assessment. Patients with diagnoses other than schizophrenia were found to have had fewer verbally aggressive

incidents after START assessment; the non-psychotic group tended to be more likely to have multiple or complex conditions. If replicated, this suggests that where assessments have to be prioritised, patients with non-psychotic illnesses or multiple diagnoses should be given priority for START assessments.

Limitations

The male-only sample precludes generalisation to female patients. Olsson et al. (2013) found that risk was more likely to decrease in female than in male psychiatric patients over time, and therefore replication with a female sample is required. Secondly, unlike Troquete et al.'s (2013) study, we had no control or comparison group. The sample size was predicated on a moderate to large effect size and it is possible that clinically meaningful and statistically significant reduction in risk behaviour might have been detected with a larger sample. In future, before services introduce risk assessment tools like START, they should consider doing so in a manner which supports research: for example, in balanced blocks so that evaluation can take advantage of the opportunity for control conditions.

References

- Abderhalden C, Needham I, Dassen T, Halfens R, Haug HJ, Fischer JE (2008). Structured risk assessment and violence in acute psychiatric wards: randomised controlled trial. *British Journal of Psychiatry* 193: 44-50. DOI: 10.1192/bjp.bp.107.045534
- Douglas KS, Skeem JL (2005). Violence risk assessment: getting specific about being dynamic. *Psychology, Public Policy, and Law* 11: 347-383. DOI: <http://dx.doi.org/10.1037/1076-8971.11.3.347>
- Foster C, Bowers L, Nijman H. (2007). Aggressive behaviour on acute psychiatric wards: prevalence, severity and management. *Journal of Advanced Nursing* 58: 140-149. DOI: 10.1111/j.1365-2648.2007.04169.x

- Gray NS, Taylor J, Snowden RJ (2011) Predicting violence using structured professional judgment in patients with different mental and behavioral disorders. *Psychiatry Research* 187: 248-253. DOI: 10.1016/j.psychres.2010.10.011
- Gunenc C, O'Shea LE, Dickens GL (2015). Prevalence and predictors of verbal aggression in a secure mental health service: Use of the HCR-20. *International Journal of Mental Health Nursing* 24: 314-323. DOI: 10.1111/inm.12130
- Hanson RK, Harris AJ (2000). Where should we intervene? Dynamic predictors of sexual offense recidivism. *Criminal Justice and Behavior* 27: 6-35.
DOI:10.1177/0093854800027001002
- Hart SD (1998). The role of psychopathy in assessing risk for violence: Conceptual and methodological issues. *Legal and Criminological Psychology* 3: 121-137. DOI: 10.1111/j.2044-8333.1998.tb00354.x
- Inett A, Wright G, Roberts L, Sheeran A (2014). Predictive validity of the START with intellectually disabled offenders. *Journal of Forensic Practice* 16: 78-88. DOI: 10.1108/JFP-12-2012-0029
- Kling RN, Yassi A, Smailes E, Lovato CY, Koehoorn M (2011). Evaluation of a violence risk assessment system (the Alert System) for reducing violence in an acute hospital: A before and after study. *International Journal of Nursing Studies*, 48: 534-539. doi: 10.1016/j.ijnurstu.2010.10.006
- Large MM, Nielssen OB (2012). Comment on risk assessment: predicting violence. *Evidence Based Mental Health*, 15: 4-4. doi:10.1136/ebmh.2011.100174
- Lipsey MW, Wilson DB (2001) *Practical meta-analysis*. London: Sage Publications.
- Nicholls TL, Gagnon N, Crocker AG, Brink J, Desmarais SL, Webster C (2007). *START Outcomes Scale (SOS)*. Vancouver, BC Mental Health & Addiction Services.

- Nonstad K, Nasset MB, Kroppan E, Pedersen TW, NØttestad JA, Almvik R, Palmstierna T (2010). Predictive validity and other psychometric properties of the Short-Term Assessment of Risk and Treatability (START) in a Norwegian high secure hospital. *International Journal of Forensic Mental Health* 9: 294-299. DOI:10.1080/14999013.2010.534958
- O'Shea, L. E., & Dickens, G. L. (2014). Short-Term Assessment of Risk and Treatability (START): Systematic review and meta-analysis. *Psychological Assessment*, 26(3), 990.
- O'Shea LE, Picchioni MM, Dickens GL (2015). The predictive validity of the Short-Term Assessment of Risk and Treatability (START) for multiple adverse outcomes in a secure psychiatric inpatient setting. *Assessment*, 22: 150-162. DOI: 10.1177/1073191115573301
- O'Shea LE, Picchioni MM, Mason FL, Sugarman PA, Dickens GL (2014a). Differential predictive validity of the Historical, Clinical and Risk Management Scales (HCR-20) for inpatient aggression. *Psychiatry Research* 220: 669-678. DOI: 10.1016/j.psychres.2014.07.080
- O'Shea LE, Dickens GL (2016) Role of assessment components and recent adverse outcomes in risk estimation and prediction: Use of the Short Term Assessment of Risk and Treatability (START) in an adult secure inpatient mental health service. *Psychiatry Research* 240: 398-405. DOI: 10.1016/j.psychres.2016.04.068
- O'Shea LE, Picchioni MM, Mason FL, Sugarman PA., Dickens GL (2014b). Predictive validity of the HCR-20 for inpatient self-harm. *Comprehensive Psychiatry* 55: 1937-1949. DOI: 10.1016/j.comppsy.2014.07.010
- O'Shea LE, Thaker DK, Picchioni MM, Mason FL, Knight C, Dickens GL (2015). Predictive validity of the HCR-20 for violent and non-violent sexual behaviour in a secure mental health service. *Criminal Behaviour and Mental Health*. doi: 10.1002/cbm.1967

- Quinn R, Miles H, Kinane C (2013). The validity of the Short-Term Assessment of Risk and Treatability (START) in a UK medium secure forensic mental health service. *International Journal of Forensic Mental Health* 12: 215-224. DOI: 10.1080/14999013.2013.832714
- Sealed Envelope Ltd. (2012). Power calculator for binary outcome superiority trial. [Online] Available from: <https://www.sealedenvelope.com/power/binary-superiority/> [Accessed Sun Jul 10 2016].
- SPSS (2009). PASW Statistics for Windows, Version 18.0. Chicago: SPSS Inc. .
- Troquete NAC, van den Brink RH, Beintema HT, Mulder TWDP, van Os RA Schoevers D, Wiersma R (2013). Risk assessment and shared care planning in out-patient forensic psychiatry: cluster randomised controlled trial. *British Journal of Psychiatry* 202: 365-371. DOI: 10.1192/bjp.bp.112.113043
- van de Sande R, Nijman HLI, Noorthoorn EO, Wierdsma, AI, Hellendoorn, E, van der Staak C, Mulder CL (2011). Aggression and seclusion on acute psychiatric wards: effect of short-term risk assessment. *British Journal of Psychiatry*, 199: 473-478. DOI: 10.1192/bjp.bp.111.095141
- Webster CD, Martin ML, Brink J, Nicholls TL, Desmarais SL (2009). Manual for the Short-Term Assessment of Risk and Treatability (START) (Version 1.1) Port Coquitlam. BC, Canada: British Columbia Mental Health and Addiction Services.
- World Health Organization. (1992). The ICD-10 classification of mental and behavioural disorders: clinical descriptions and diagnostic guidelines.

Table 1: Distribution of specific risk estimate (SRE) scores according to the Short-Term Assessment of Risk and Treatability (START) scale

Outcome	Specific Risk Estimate			
	Low (%)	Moderate (%)	High (%)	Missing (%)
Violence	29 (58)	8 (16)	7 (14)	6 (12)
Self-harm	31 (62)	6 (12)	0 (0)	13 (26)
Suicide	27 (54)	2 (4)	0 (0)	21 (42)
Unauthorised leave	28 (56)	9 (18)	1 (2)	12 (24)
Substance abuse	31 (62)	3 (6)	0 (0)	16 (32)
Self-neglect	24 (48)	16 (32)	3 (6)	7 (14)
Victimisation	30 (60)	9 (18)	3 (6)	8 (16)

Table 2: Total number of incidents in each measured category and mean number of incidents per patient three months before and three months after assessment.

Outcome	Before Assessment	Mean (SD)	After Assessment	Mean (SD)	Statistical Test
Physical aggression	141	2.82 (5.98)	87	1.74 (4.31)	$t[50] = -0.63, p=0.53$
Verbal aggression	244	4.88 (7.17)	168	3.36 (5.67)	$t[50] = 1.46, p=0.15$
Self-harm	13	0.26 (0.94)	9	0.18 (0.44)	$t[50] = -1.07, p=0.29$
Victimisation	17	0.34 (0.69)	12	0.24 (0.62)	$t[50] = -0.92, p=0.36$
Self-neglect	3	0.06 (0.24)	7	0.14 (0.53)	$t[69] = 1.15, p=0.25$
Substance abuse	3	0.06 (0.42)	1	0.02 (0.14)	$t[50] = -1.32, p=0.19$
Unauthorised leave	2	0.04 (0.28)	3	0.06 (0.24)	$t[50] = 0.62, p=0.54$

Table 3: Number of patients with a recorded incident in each measured category three months before and three months after assessment.

Outcome	Before Assessment (%)	After Assessment (%)	Statistical Test
Physical Aggression	28 (56)	23 (46)	$\chi^2[1]=0.640, p=0.424$
Verbal Aggression	32 (64)	27 (54)	$\chi^2[1]=0.661, p=0.416$
Self-Harm	5 (10)	8 (16)	$\chi^2[1]=0.354, p=0.552$
Victimisation	12 (24)	8 (16)	$\chi^2[1]=2.080, p=0.149$
Self-Neglect^a	3 (6)	4 (8)	-
Substance Abuse^a	1 (2)	1 (2)	-
Unauthorised Leave^a	1 (2)	3 (6)	-

^a N.B. Expected cell frequency insufficiently large for analysis